Variational Data Assimilation of Microwave Radiobrightness Observations for Land Surface Hydrology Applications

Rolf H. Reichle, Dennis B. McLaughlin, and Dara Entekhabi

Abstract—Our ability to accurately describe large-scale variations in soil moisture is severely restricted by process uncertainty and the limited availability of appropriate soil moisture data. Remotely sensed microwave radiobrightness observations can cover large scales but have limited resolution and are only indirectly related to the hydrologic variables of interest. We describe a four-dimensional (4-D) variational assimilation algorithm that makes best use of available information while accounting for both measurement and model uncertainty. The representer method used here is more efficient than a Kalman filter because it avoids explicit propagation of state error covariances. In a synthetic example, which is based on a field experiment, we demonstrate estimation performance by examining data residuals. Such tests provide a convenient way to check the statistical assumptions of the approach and to assess its operational feasibility. Internally computed covariances show that the estimation error decreases with increasing soil moisture. An adjoint analysis reveals that trends in model errors in the soil moisture equation can be estimated from daily L-band brightness measurements, whereas model errors in the soil and canopy temperature equations cannot be adequately retrieved from daily data alone. Nonetheless, state estimates obtained from the assimilation algorithm improve significantly on prior model predictions derived without assimilation of radiobrightness data.

Index Terms—Data assimilation, land surface hydrology, representer method, soil moisture.

I. INTRODUCTION

SOIL moisture at the land surface is a key variable for weather and climate prediction, flood forecasting, and the determination of groundwater recharge. Exfiltration and infiltration fronts are partially controlled by the state of the soil moisture profile. In addition, diurnal variations in land surface fluxes and states are affected by conditions below the surface. Although soil moisture is of great importance, it is difficult to monitor routinely, especially over regional or continental scales. *In situ* point observations are problematic for characterizing large-scale variations, and remotely sensed passive microwave measurements can only be indirectly related to moisture in the 5 cm layer nearest to the surface.

R. H. Reichle is with the Goddard Earth Sciences and Technology Center, University of Maryland Baltimore County, Baltimore, MD 21250 USA (e-mail: reichle@janus.gsfc.nasa.gov).

D. B. McLaughlin and D. Entekhabi are with the Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

In this paper, we describe a method for deriving dynamically consistent estimates of the soil moisture profile by assimilating remotely sensed passive microwave measurements into a land surface model. The primary model states are soil moisture and soil temperature and the measured variable is the L-band (1.4 GHz) brightness temperature. The state and measurement equations include additive errors, which are assumed to be random variables with specified statistical properties. The estimation process can be reduced to the solution of a constrained least-squares problem, which is difficult to solve because the number of unknowns can be very large and the constraining model is nonlinear.

Most soil moisture assimilation studies have focused on estimation problems with only one spatial (vertical) dimension, e.g., [1]–[5]. Since in one dimension, the state vector is small, these studies can apply advanced sequential or variational assimilation techniques that account for model errors and error covariance propagation. Unfortunately, such advanced methods do not scale well with the size of the problem. Recently, Houser *et al.* [6] investigated four-dimensional (4-D) soil moisture assimilation using *in situ* and remote sensing observations, but their assimilation methods neglect model errors and error covariance propagation. For a detailed review of soil moisture assimilation, see [7]. Our goal in this paper is to develop a 4-D land data assimilation algorithm that accounts for model errors and fully incorporates process dynamics into the estimates.

4-D variational data assimilation methods have been widely used in meteorological and oceanographic applications [8]–[10], although relatively few of these have accounted for model error. The variational approach offers the possibility of achieving the accuracy of advanced sequential estimation algorithms such as the Kalman filter [11] with the computational efficiency of suboptimal methods. This is possible because variational methods do not explicitly evaluate the large error covariance matrices which are propagated by Kalman filtering algorithms. Moreover, variational methods are particularly well suited for the highly nonlinear soil moisture problem. Various computational enhancements developed in recent years have greatly improved the practical potential of variational data assimilation [12]. These include the indirect iterated representer method [7], [12], [13], which we use in our soil moisture estimation algorithm.

We begin in Sections II–IV with brief reviews of the assimilation method and the land surface model used in our soil moisture application. In Section V, we describe the setup of a synthetic experiment used to demonstrate the performance of the algo-

rithm. Section VI discusses the results of this experiment. In the final section, we summarize the implications of the synthetic experiment and assess the prospects for variational assimilation of soil moisture on an operational basis.

II. VARIATIONAL DATA ASSIMILATION

Model predictions and measurements both provide useful information about the actual state of the soil. The objective of our soil moisture data assimilation algorithm is to combine these two sources of information. The relevant models for the soil moisture application are based on mass and energy balance equations that describe the temporal and spatial evolution of soil moisture, soil temperature, and canopy temperature. These variables may be approximated by a set of spatially discretized time-dependent states defined at the cells of a three-dimensional (3-D) computational grid.

If all of the spatially discretized states are assembled in a time-dependent N_Y-dimensional vector $Y(t)$ the hydrologic model can be expressed as

$$\frac{dY}{dt} = \varphi(Y) + \omega \quad Y|_{t=0} = Y_0(\beta). \tag{1}$$

The N_Y-dimensional vector ω is composed of a set of model errors that are treated as random forcing terms in the state equation. The initial condition is parameterized with a random vector β, which may have a dimension less than N_Y. These random variables are defined in terms of their first- and second-order moments. In particular, the specified prior mean values $\overline{\omega}$ and $\overline{\beta}$ and prior covariances C_ω and C_β summarize our information about the errors prior to data assimilation. For convenience, we assume that $\overline{\omega}$ is zero. This assumption could be relaxed if evidence suggested the presence of a systematic bias (nonzero mean) in the model error. Note that temporal correlations in the model errors can lead to a bias in the state even when the model error itself is unbiased. Deterministic forcings such as observed meteorological inputs are included in the nonlinear operator φ.

The radiobrightness measurements to be assimilated by the algorithm can be related to the moisture and temperature states with a nonlinear radiative transfer model. It is convenient to assemble all brightness temperature measurements available over the temporal "assimilation interval" $t \in [0, t_f]$ in an N_Z-dimensional vector Z. The kth element Z_k of Z is the scalar measurement obtained at time t_k over an area centered on location x_k. Next, let $m_k(Y(t_k))$ denote the true radiobrightness at location x_k and time t_k, where $m_k(\cdot)$ is derived from a radiative transfer model (Section IV-A). The complete measurement vector can then be written as an integral over a set of Dirac delta functions that sample the true radiobrightness values at all of the measurement times and locations

$$Z = M[Y] + v \quad M[Y] \equiv \int_0^{t_f} D(t)m(Y(t))dt. \tag{2}$$

Here, $D(t)$ is a diagonal matrix of dimension N_Z by N_Z with the kth diagonal element equal to $\delta(t - t_k)$. Moreover, $m(Y(t))$ and v are N_Z-dimensional vectors with kth components m_k and v_k, respectively. The random vector v in (2) accounts for additive measurement errors and is defined in terms of its first and second moments. We assume that v is zero mean with a specified covariance matrix C_v of dimension N_Z by N_Z.

The data assimilation algorithm is designed to balance the effects of uncertain model errors, initial condition parameters, and measurement errors. This is achieved by identifying the model errors and initial condition parameters which minimize a weighted least-squares (minimum variance Bayesian) objective function J

$$\begin{aligned} J =& (Z - M[Y])^{\mathrm{T}} C_v^{-1} (Z - M[Y]) \\ &+ (\beta - \overline{\beta})^{\mathrm{T}} C_\beta^{-1} (\beta - \overline{\beta}) \\ &+ \int_0^{t_f} \int_0^{t_f} \omega(t')^{\mathrm{T}} C_\omega^{-1}(t', t'') \omega(t'') dt' dt'' \\ &+ 2 \int_0^{t_f} \lambda^{\mathrm{T}} \left(\frac{dY}{dt} - \varphi(Y) - \omega \right) dt. \end{aligned} \tag{3}$$

This objective function minimizes the aggregate error over the assimilation interval. The first term accounts for the misfit between the data vector Z and the measurement predictions $M[Y]$, normalized by the measurement error covariance C_v. The second and third terms penalize normalized deviations of the uncertain inputs from specified prior mean values. The final term is obtained by adjoining the state equation to the performance index with the time-dependent N_Y-dimensional Lagrange multiplier vector λ. This constraint ensures that the state estimates produced by the assimilation procedure satisfy the hydrologic model within the range specified by the model error statistics.

Estimates of the parameter vector $\hat{\beta}$ and the model error $\hat{\omega}(t)$ are obtained by setting the first variation of the objective function (3) equal to zero. This yields a set of so-called Euler-Lagrange equations, which constitute a two-point boundary value problem [7]

$$\frac{d\hat{Y}}{dt} = \varphi(\hat{Y}) + \hat{\omega} \quad \hat{Y}|_{t=0} = Y_0(\hat{\beta}) \tag{4}$$

$$-\frac{d\lambda}{dt} = \frac{\partial \varphi}{\partial Y}\bigg|_{\hat{Y}}^{\mathrm{T}} \lambda + \frac{\partial m}{\partial Y}\bigg|_{\hat{Y}}^{\mathrm{T}} D(t) C_v^{-1} (Z - M[\hat{Y}]) \quad \lambda|_{t=t_f} = 0 \tag{5}$$

$$\hat{\omega} = \int_0^{t_f} C_\omega(t, t') \lambda(t') dt' \quad \hat{\beta} = \overline{\beta} + C_\beta \frac{\partial Y_0}{\partial \beta}\bigg|_{\hat{\beta}}^{\mathrm{T}} \lambda|_{t=0} \tag{6}$$

where $\partial \varphi / \partial Y$ and $\partial m / \partial Y$ are N_Y by N_Y and N_Z by N_Y matrices, respectively. Here and in the following, the estimates are denoted with a hat.

The first, or forward, Euler–Lagrange (4) ensures that the estimates obey the state (1). The second, or adjoint (5), describes the temporal evolution of the adjoint variable (or Lagrange multiplier) λ. This variable is closely related to the model error and parameter estimates, which are obtained from (6). Note that the adjoint (5) is forced by differences between measurements and the corresponding model predictions. This equation is integrated backward in time, starting from a homogeneous terminal condition. The model error estimate is the convolution of the adjoint variable over the model error covariance C_ω, while the initial condition parameter estimate is a weighted sum of the adjoint

values at the initial time, with weights derived from the covariance matrix C_β. Once the unknown model errors and parameters are estimated, the forward equation may be used to derive the state estimate $\hat{Y}$ at any time or location. Note that the adjoint variable λ contains all of the measurement information needed to derive the updated estimates $\hat{\omega}$, $\hat{\beta}$, and $\hat{Y}$. These estimates can be obtained without explicit computation of the state estimation error covariance matrix.

III. THE REPRESENTER ALGORITHM

The coupled set of nonlinear Euler–Lagrange equations described in the previous section must be solved with an iterative numerical algorithm. Although several possibilities are available for solving such two-point boundary value problems, we have found that the best combination of efficiency and performance is provided by the indirect iterated representer algorithm [12]. This algorithm is particularly efficient because it propagates covariance information implicitly (through the Euler–Lagrange equations) rather than explicitly (with a large covariance matrix). As a result, the method is able to provide statistically optimal estimates without the simplifications that have been used in other large-scale soil moisture estimation applications. Fig. 1 summarizes the algorithm.

In the iterated representer algorithm, the estimates are derived by solving a sequence of linear estimation problems until convergence is achieved. Although there is no proof or guarantee that this sequence converges, experience has shown that it converges in practice for models that are reasonably close to linear. At iteration level $\eta + 1$, we linearize the model and measurement equations around the nominal trajectory Y^η and around β^η, which are the estimates from the previous iteration η

$$
\frac{dY}{dt} = \varphi(Y^\eta) + \left.\frac{\partial\varphi}{\partial Y}\right|_\eta (Y - Y^\eta) + \omega
$$

$$
Y|_{t=0} = Y_0(\beta^\eta) + \left.\frac{\partial Y_0}{\partial\beta}\right|_\eta (\beta - \beta^\eta) \tag{7}
$$

$$
Z = M[Y^\eta] + L[Y - Y^\eta] + v
$$

$$
L[Y] \equiv \int_0^{t_f} D(t)\left.\frac{\partial m}{\partial Y}\right|_\eta Y(t)dt. \tag{8}
$$

The linear functional $L[\cdot]$ can be viewed as the derivative (or slope) of the measurement operator $M[\cdot]$. The Euler–Lagrange equations for the linearized problem are

$$
-\frac{d\lambda^{\eta+1}}{dt} = \left.\frac{\partial\varphi}{\partial Y}\right|_\eta^{\mathrm{T}} \lambda^{\eta+1} + \left.\frac{\partial m}{\partial Y}\right|_\eta^{\mathrm{T}} D(t)b \qquad \lambda|_{t=t_f}^{\eta+1} = 0 \tag{9}
$$

$$
\omega^{\eta+1} = \int_0^{t_f} C_\omega(t,t')\lambda^{\eta+1}(t')dt'
$$

$$
\beta^{\eta+1} = \overline{\beta} + C_\beta \left.\frac{\partial Y_0}{\partial\beta}\right|_\eta^{\mathrm{T}} \lambda|_{t=0}^{\eta+1} \tag{10}
$$

$$
\frac{dY^{\eta+1}}{dt} = \varphi(Y^\eta) + \left.\frac{\partial\varphi}{\partial Y}\right|_\eta (Y^{\eta+1} - Y^\eta) + \omega^{\eta+1}
$$

$$
Y|_{t=0}^{\eta+1} = Y_0(\beta^\eta) + \left.\frac{\partial Y_0}{\partial\beta}\right|_\eta (\beta^{\eta+1} - \beta^\eta) \tag{11}
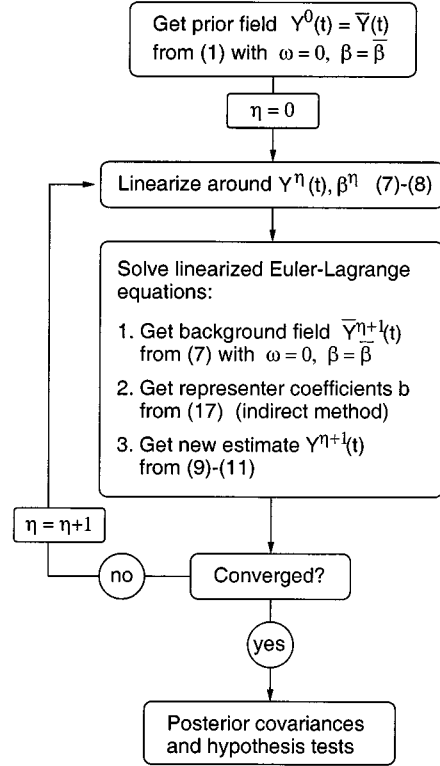$$

Fig. 1.   Flowchart for the indirect iterated representer method.

where in (9) we defined

$$
b \equiv C_v^{-1}\left(Z - M[Y^\eta] - L\left[Y^{\eta+1} - Y^\eta\right]\right). \tag{12}
$$

Aside from the factor C_v^{-1}, the coefficient vector b is equal to the posterior data misfit. Note that the linearized Euler–Lagrange equations are coupled through the vector b (that is, b depends on $Y^{\eta+1}$, which depends on b). This makes the problem difficult to solve. The idea of the representer algorithm is to find b independently and thus break the coupling of the Euler–Lagrange equations. Then (9) through (11) could be solved in sequence, starting with the adjoint equation and progressing through the update equations and the linearized state equation giving (in order) $\lambda^{\eta+1}$, $\omega^{\eta+1}$, $\beta^{\eta+1}$, and $Y^{\eta+1}$.

The decoupling of the linear Euler–Lagrange equations relies on a series expansion of the estimated fields

$$
Y^{\eta+1}(t) = \overline{Y}^{\eta+1}(t) + \sum_{k=1}^{N_Z} b_k \Upsilon^k(t) \tag{13}
$$

$$
\lambda^{\eta+1}(t) = \sum_{k=1}^{N_Z} b_k \Lambda^k(t). \tag{14}
$$

The leading term of the state representer expansion (13) is the background field $\overline{Y}^{\eta+1}$, which is obtained from (7) using $\omega = \overline{\omega} \equiv 0$ and $\beta = \overline{\beta}$. Note that the background field $\overline{Y}^{\eta+1}$ is an approximation of the prior field $\overline{Y}$, which is derived from (1) with $\omega = \overline{\omega} \equiv 0$ and $\beta = \overline{\beta}$. The N_Z coefficients b_k, as defined in (12), are called the representer coefficients. The same coefficients are used in (13) and (14).

The N_Z basis functions Υ^k of the expansion, called state representer functions, are the linearized prior cross-covariances between the (scalar) measurement predictions $M_k[Y]$ and the state Y [7]. The N_Z functions Λ^k of (14) are called the adjoint representers. In each case, the superscript k stands for the number of the measurement corresponding to the representer function in question. Each Υ^k and each Λ^k is an N_Y-dimensional vector defined over $[0, t_f]$. The representer functions are derived from the Euler–Lagrange equations rather than specified *a priori*. Inserting (13) and (14) into (9), (10), and (11), we obtain

$$-\frac{d\Lambda^k}{dt} = \frac{\partial\varphi}{\partial Y}\bigg|_{\eta}^{\mathrm{T}}\Lambda^k + \frac{\partial m_k}{\partial Y}\bigg|_{\eta}^{\mathrm{T}}\delta(t - t_k) \qquad \Lambda^k|_{t=t_f} = 0 \tag{15}$$

$$\frac{d\Upsilon^k}{dt} = \frac{\partial\varphi}{\partial Y}\bigg|_{\eta}\Upsilon^k + \int_0^{t_f} C_\omega(t, t')\Lambda^k(t')dt'$$

$$\Upsilon^k|_{t=0} = \frac{\partial Y_0}{\partial\beta}\bigg|_{\eta} C_\beta \frac{\partial Y_0}{\partial\beta}\bigg|_{\eta}^{\mathrm{T}}\Lambda^k|_{t=0}. \tag{16}$$

Whereas the adjoint (9) is forced with a linear combination of the posterior data misfit (or b), the adjoint representer (15) is forced with a single impulse, which corresponds to one measurement time and location and is independent of the current estimate $Y^{\eta+1}$. We can therefore solve for the representer fields without knowing the current estimate, which is consistent with the interpretation of the representers as prior covariances. Strictly speaking, however, the linearization around the previous estimate Y^η implies that the representer fields depend on the data after the first iteration. Note also that the representer fields and coefficients change at each iteration.

The representer expansions enable us to derive a representer coefficient expression which does not depend on $Y^{\eta+1}$. Inserting (13) and (14) in (12), we find

$$b = (R + C_v)^{-1}\left(Z - M[Y^\eta] - L\left[\overline{Y}^{\eta+1} - Y^\eta\right]\right). \tag{17}$$

The representer matrix $R_{kl} \equiv L_k[\Upsilon^l]$ with $k, l = 1 \ldots N_Z$ is the (prior) covariance matrix of the measurement predictions $M_k[Y]$. Since (17) is in terms of the prior data misfit, we could now compute the estimates by explicitly calculating the representer functions Υ^k and the representer matrix R and then solving (17) for the representer coefficients b. This approach is called the direct iterated representer method. There is, however, a faster way to get the estimates, which also avoids the computation and storage of the representer matrix. This technique is called the indirect iterated representer method. Rather than computing all N_Z individual representer fields Υ^k, we calculate only the linear combinations of representer functions that are needed in a conjugate gradient solver of (17). For details, see [7], [12]. The computational savings are discussed in Section VI-E.

A. Posterior Covariances

After the iteration has converged, we can opt for calculating the posterior covariances of the estimates. For large problems, however, we can at best compute the posterior covariance of the

measured variables. The computational effort for the state error covariances is simply too large. Moreover, the problem at hand is nonlinear and posterior covariances based on the linearizations (7) and (8) can only approximate the true covariances.

Since $v = Z - M[Y]$ is the true measurement error (2), we define $\tilde{v} \equiv M[Y] - M[\hat{Y}] \approx L[Y - \hat{Y}]$ as the (posterior) estimation error of the measurement predictions and $\hat{v} \equiv Z - M[\hat{Y}]$ as the estimate of the measurement error, or equivalently, the vector of posterior data residuals. The corresponding posterior covariances are [7], [9]

$$C_{\hat{v}} = C_v(R + C_v)^{-1}C_v \qquad C_{\tilde{v}} = C_v - C_{\hat{v}}. \tag{18}$$

Recall that the representer matrix R is the prior covariance of the measurement predictions. We can use $C_{\hat{v}}$ to normalize the posterior data residuals and to test whether their distribution is standard normal and uncorrelated in space and time (Section VI-B). This test is particularly powerful when actual data are assimilated, because it provides a way to check whether the estimator is operating in accordance with its statistical assumptions. It must be stressed, however, that this test relies on the linearizations that we have adopted.

IV. LAND SURFACE MODEL FOR DATA ASSIMILATION

The models that form the basis for soil moisture data assimilation must capture key physical processes while remaining efficient enough to make large-scale estimation computationally feasible. These are potentially conflicting requirements that must be traded off when a model is selected. For the variational approach, we also require a numerically well behaved adjoint model. The development of such an adjoint is a time-consuming task and there are currently no adjoint models available for the commonly used land surface schemes. We have developed a hydrologic model which meets the requirements of variational data assimilation. The model is described in [14] and in detail in [7]. Its most important features are summarized in this section.

A. Vertical Moisture and Energy Transport

Our model of coupled moisture and heat transport is a soil–vegetation–atmosphere transfer scheme (SVAT) that relies on many of the concepts used in other hydrologic modeling applications. Vertical soil moisture transport in each pixel is governed by Richards' equation, with soil hydraulic properties described by relationships from [15]. We use five subsurface layers in the Richards' equation discretization, located at 0–5 cm, 5–15 cm, 15–30 cm, 30–55 cm, and 55–90 cm. The lower boundary condition for the soil moisture equation is defined by gravity drainage (zero pressure gradient).

In order to achieve maximum computational efficiency, vertical energy transport in each pixel is described with a one-layer force-restore method [16] rather than the full heat equation. This approach is ideal for our remote sensing application since it focuses on diurnal soil temperature fluctuations in the upper layer, which effects microwave radiation. The deep soil temperature is specified as the seasonal average of the observed air temperature. Vertical moisture and heat fluxes through the vegetation layer are modeled with a resistance network approach [17].

At the land surface, we assume near-neutral atmospheric conditions. Finally, the brightness temperature is related to the land surface states with a grey body radiative transfer model [18].

B. Horizontal Variability

In order to be operationally useful, our data assimilation technique must consider horizontal as well as vertical variability. Fortunately, fluxes and length scales in the vertical and horizontal dimensions are much different. This scale disparity can be exploited to improve computational efficiency without any significant sacrifice in accuracy. In particular, we assume that horizontal moisture and heat fluxes in the unsaturated zone are negligible. As a result, horizontal variations in our soil moisture estimates reflect spatial correlation in meteorological inputs, land cover, and soil texture rather than lateral transport. This assumption is reasonable for terrain with moderate relief over the spatial scales under consideration here (Fig. 2). It enables us to break the model domain into a grid of one-dimensional (1-D) vertical cells or pixels. Our 4-D assimilation algorithm accounts for horizontal correlations in the forcing errors as well as for unmodeled horizontal fluxes through the model error terms. The assimilation algorithm derives estimates of the soil moisture in the five layers associated with each pixel and estimates of the surface soil temperature and canopy temperature of each pixel.

Satellite observations of L-band brightness temperatures will likely be available only over spatial scales much larger than the scales of the meteorological data and other inputs. This requires us to downscale from coarse measurements to higher resolution estimates. Downscaling can be carried out in a straightforward manner within the framework of our algorithm [14]. In this paper, we limit ourselves to estimating land surface states at the scale of the brightness observations.

V. A Synthetic Experiment

In the following sections, we investigate the performance of our assimilation algorithm using an experiment based on synthetically generated radiobrightness data. Synthetic experiments are ideally suited for algorithm performance tests since all of the uncertain inputs are known by design. In particular, synthetic experiments allow us to isolate the effects of the nonlinearities in the hydrological model on the quality of the estimates. Such experiments are an indispensable first step toward a field application, although they clearly cannot replace tests based on actual observations. Recognizing these limitations, it is useful to examine the results from well-controlled synthetic experiments.

The synthetic test problem considered in this paper is based on the Southern Great Plains (SGP97) hydrology experiment in Central Oklahoma [19]. The 80 km $\times$ 160 km area chosen for the synthetic experiment is shown in Fig. 2. We divide this area into 16 by 32 pixels of 5 km $\times$ 5 km. Soil temperature and soil moisture are estimated in each of these 512 pixels. Fig. 2 also displays land cover data (from the SGP97 Data Archive) and soil texture data (from the Earth System Science Center, Pennsylvania State University, State College). The meteorological inputs are taken from the Oklahoma Mesonet database and interpolated to the model grid with inverse-square distance weigths.
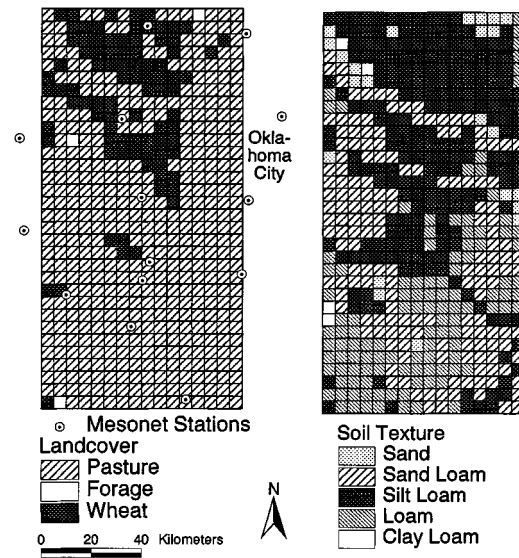

Fig. 2. Area for the synthetic experiment. The horizontal grid of 16×32 pixels (80 km $\times$ 160 km) is shown together with (left) the locations of the Oklahoma Mesonet stations and the land cover classes and (right) with the soil texture classes (from [14]).

Our experiment extends for eight days from June 27, 1997 to July 4, 1997, which are covered by a single assimilation interval. Soil moisture estimates are derived at a time step of 30 min.

The radiobrightness measurements obtained for our experiment are based on synthetically generated "true" initial condition parameters β and time-dependent model error fields ω. These are used in the forward model (1), which is solved to give a set of "true" system states. Recall that the "prior" state is the solution to (1) when the initial conditions and model errors are set to their prior values. This prior state can be viewed as the "best guess" of the true states available without the benefit of radiobrightness measurements. Fig. 3 shows the true and prior top node saturation for our experiment at four times during the assimilation interval. We define soil saturation to vary between zero and one.

The "true" brightness temperatures are obtained by using the "true" states as inputs to the radiative transfer model. The vector Z of synthetic brightness measurements is obtained by adding random measurement errors to the output of this model. Synthetic brightness measurements are generated once daily at every pixel in the model domain at 10:00 AM local time, which is typical of the SGP97 field experiment. This yields seven images and a total of $N_Z = 7 \cdot 16 \cdot 32 = 3584$ scalar measurements. The assimilation algorithm estimates the land surface states from the noisy data Z and the prior solution.

The spatial and temporal correlation functions of the uncertain inputs that are needed for the generation of the synthetic true fields and for the estimation algorithm depend on the given model and field setting. These statistics are very difficult to determine, and their characterization is beyond the scope of this paper. Here, we only aim to prove the concept of soil moisture assimilation with synthetic experiments. This does not critically depend on the exact shapes and scales of the correlation functions, and we specify conditions that in our experience are appropriate for the experiment area and our model.
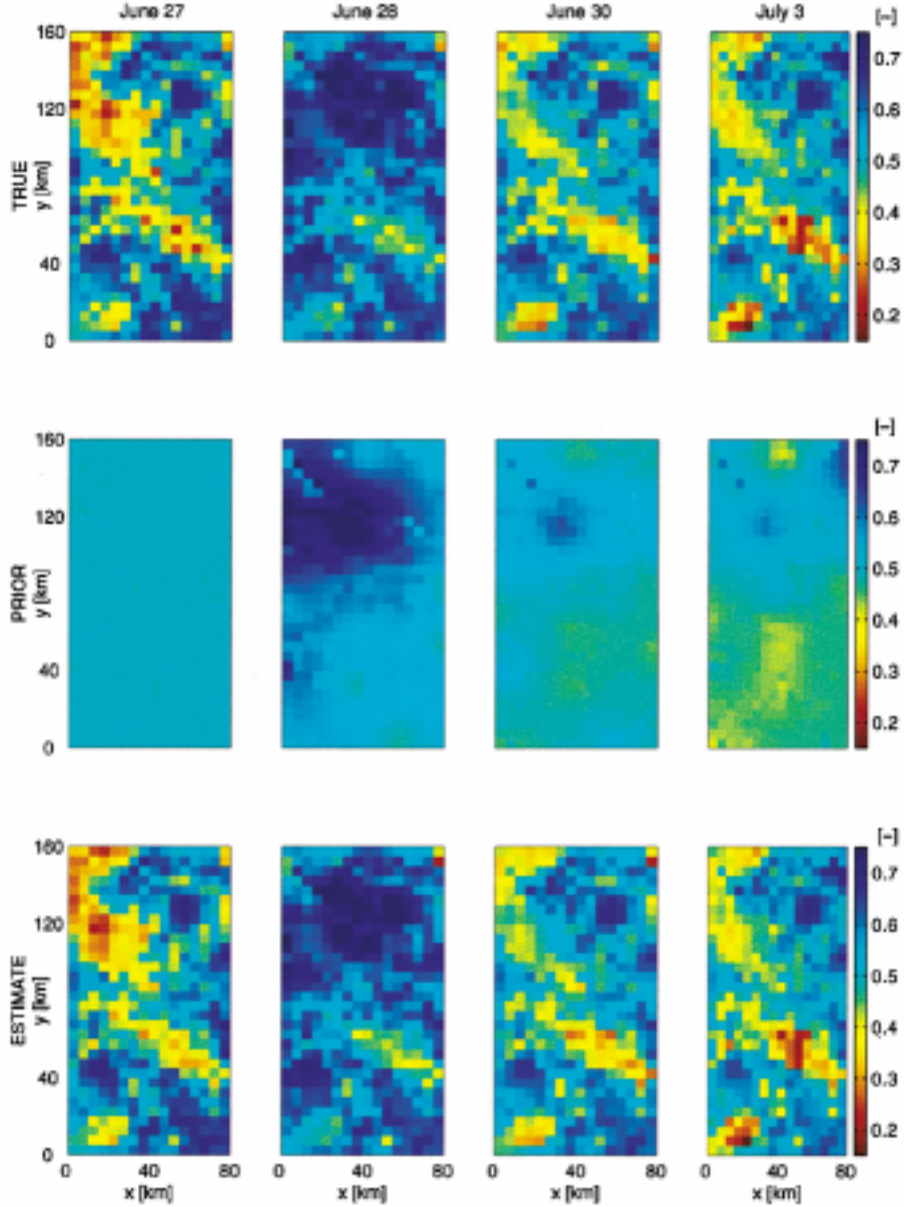
Fig. 3. First row shows the true top node saturation on four days at 10:00 AM. The second and third rows show the prior and the estimated top node saturation at the same times, respectively.

The initial conditions and model errors generated in our synthetic experiment are zero-mean normally distributed random fields with exponential spatial and temporal correlation functions. The only uncertain initial condition parameter in each pixel is the total initial water storage across the modeled profile. The initial saturation profile is related to this parameter through a shape function which can be specified arbitrarily. Here we use a hydrostatic shape. In an operational setting, the shape could be derived from the estimated profile at the final time of a preceding assimilation interval. The standard deviation of the initial top node saturation is 0.12 and the correlation length is 50 km. The initial upper layer soil temperature is set equal to the initial air temperature and is assumed to be known perfectly (its memory is only a few hours and has little impact on the estimates).

Model errors are represented as unknown fluxes in the near-surface soil moisture, soil energy, and canopy energy balance equations. Each of these errors is zero mean and has a standard deviation equivalent to 50 W m^{-2}. The model error correlation lengths are all 6 km, and their correlation times are all 10 h. The random measurement errors added to the brightness temperature values are spatially and temporally uncorrelated with a standard deviation of 5 K. The absence of spatial correlation of the measurement error is not a constraint imposed by the algorithm but is a simplification adopted for convenience.

VI. Results and Discussion

A. Soil Moisture Estimates

We begin the discussion of our synthetic data assimilation experiment with a review of the soil moisture saturation estimates produced by the algorithm. Fig. 3 shows the true, the prior, and the estimated top node saturation at four times during the

one-week assimilation interval. The temporal evolution of the top node saturation is governed by the initial condition (which tends to persist during drydown periods), precipitation, soil texture, and land cover. The value of the observations for retrieving the true soil moisture is obvious. In the absence of brightness data, our best guess for the top node saturation is the prior solution shown in the middle row of Fig. 3, which is quite far from the true saturation. Once we assimilate the brightness data, we are able to accurately estimate the initial condition and the evolution of the top node soil moisture.

The profile estimates obtained from this particular synthetic experiment are as good as the surface estimates (not shown). In practice, the feasibility of retrieving subsurface moisture profiles from surface measurements depends on the accuracy and the physical realism of the land surface model and the associated error statistics. Since the subsurface states cannot be remotely sensed at the pixel scale they can only be estimated by using the hydrologic model to propagate information downward from the surface. The time required for such propagation is significantly longer than the length of a single assimilation interval. This implies that we need to conduct a multi-interval test over several months before we can fully demonstrate the feasibility of estimating subsurface profiles from surface measurements. Such a test requires a procedure for reinitializing multiple intervals. Reinitialization is beyond the scope of this paper but is discussed in [14].

B. Posterior Covariances and Data Residuals

In our synthetic experiment, we can assess the accuracy of the estimates produced by the data assimilation algorithm simply by comparing the estimates to the synthetic true values, as we have done in Fig. 3. In practice, we do not know the true states and we must assess performance in other ways. One of the best options for the soil moisture application is to examine the differences between measured and estimated radiobrightness values. These are the posterior data residuals $\hat{v} \equiv Z - M[\hat{Y}]$ discussed in Section III-A. For the performance analysis, it is convenient to construct a set of normalized posterior data residuals $C_{\hat{v}}^{-1/2}\hat{v}$. If the estimation problem were linear, these normalized posterior residuals would have a unit normal distribution $N(0,1)$ (since all input errors are normally distributed), and they would be spatially and temporally uncorrelated. If the normalized residuals satisfy these properties at least approximately for our nonlinear problem, when all covariances are based on linearizations about the final estimate, we can assume with reasonable confidence that the estimation process is near-optimal.

Before we can normalize the posterior residuals, we must compute the theoretical prior and posterior covariances of the brightness temperature at the measurement times and locations (Section III-A). Fig. 4 shows the theoretical area-average estimation error standard deviation for each of the seven synthetic brightness images before and after assimilation, derived from R and $C_{\hat{v}}$, respectively (we define the area-average standard deviation as the square root of the area-average variance). The spatial and temporal average of the prior error standard deviation is about 16 K, with variations from a few K up to 40 K within an image. When we assimilate the brightness images, which have a constant measurement error standard deviation of 5 K,
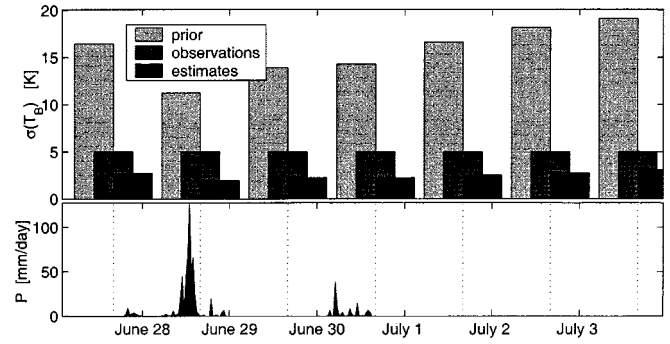

Fig. 4.   Upper panel shows the area-average error standard deviation of the brightness temperature $\sigma(T_B)$. Each group of bars refers to one observation time, which is indicated with vertical dotted lines in the lower panel. The first bar of each group (light gray) shows $\sigma(T_B)$ of the prior solution, the second bar (dark gray) shows $\sigma(T_B)$ of the observations, and the third bar (black) shows $\sigma(T_B)$ of the estimates. The lower panel shows a time series of the area-average precipitation.

TABLE I
STATISTICS OF THE NORMALIZED POSTERIOR DATA RESIDUALS

Observation time	Mean with 95% confidence interval [–]	K-S test for $N(0,1)$ at $\alpha = .05$	Variance with 95% confidence interval [–]	
1	0.012±0.087	accept	1.004	[0.89,1.14]
2	0.026±0.086	accept	0.990	[0.88,1.12]
3	-0.012±0.089	accept	1.039	[0.92,1.18]
4	-0.039±0.088	accept	1.019	[0.90,1.16]
5	-0.073±0.087	reject	1.005	[0.89,1.14]
6	-0.097±0.086	accept	0.975	[0.87,1.11]
7	-0.074±0.087	accept	1.004	[0.89,1.14]
all	-0.037±0.033	reject	1.005	[0.96,1.05]

we achieve an estimation error standard deviation of about 2.5 K, which is relatively homogeneous across the domain.

In Fig. 4, we can also see the influence of precipitation events on the uncertainty in the brightness estimates (and hence in soil moisture). The prior and posterior error standard deviations decrease after the rain event and then slowly increase again as the soil dries out. Generally, we find that the estimation error standard deviation decreases with increasing soil moisture. In other words, precipitation events tend to wipe out the memory of the system, including the effect of model errors.

We can now turn to the discussion of the posterior data residuals $\hat{v}$. Although their covariance $C_{\hat{v}}$ is not strictly diagonal, it is very diagonally dominant. For reasons of numerical accuracy (the computed $C_{\hat{v}}$ is not perfectly symmetric), we neglect off-diagonal elements of $C_{\hat{v}}$ and simply normalize the posterior residuals with their theoretical standard deviation. Table I shows sample mean values and variances of the normalized posterior data residuals for each radiobrightness image, together with 95% confidence intervals. We find that the sample means for six of the seven residual images are compatible with the zero mean hypothesis at a 5% significance level. For the sixth image and the sample of all residuals, the 95% confidence interval does

TABLE II
OBJECTIVE FUNCTION AND NUMBER OF MODEL INTEGRATIONS

Iteration	prior	1	2	3	4
Objective (J) [–]	25,098	3,424	3,394	3,461	3,485
Model integrations	–	603	191	243	163

not include zero and we have a slight overall negative bias. A Kolmogorov-Smirnov [20] test indicates that the hypothesis of a standard normal distribution cannot be rejected at a 5% significance level for six of the seven images (Table I). The normality hypothesis is rejected for the fifth image as well as for the sample of all residuals. Finally, the sample variance of all normalized posterior residuals is statistically indistinguishable from one.

If we plot the normalized posterior residuals (not shown), we find that they exhibit very little structure in either time or space, confirming the theoretical prediction that they should be nearly uncorrelated. A quantitative investigation of the two-point autocorrelation function corroborates this result. In summary, the normalized posterior data residuals do not fulfill all the criteria associated with an optimal linear estimator but they share many of the same qualitative features. Given the nonlinearities involved in the assimilation, there is no reason to expect the posterior residuals to be either strictly normal or uncorrelated. However, the general symmetry of the probability distribution and the weak correlation of the posterior residuals suggest that the estimator is acting nearly optimally.

C. Reduced Objective Function

The value of the objective function (3) provides another useful indicator of estimator performance. If the assumed statistical properties of the measurement, model, and parameter errors are correct this objective should be approximately χ^2-distributed with a mean of $\overline{J} = N_Z$ and a standard deviation of $\sigma_J = \sqrt{2N_Z}$ [9]. For our test problem the theoretical values are $\overline{J} = 3584$ and $\sigma_J = 85$. Table II shows the prior and reduced values of the objective as a function of the iteration. In the first iteration, the objective function decreases significantly from its prior value of 25 098 to 3424. But the objective function value does not continue to decrease monotonically, as it would in a steepest descent search. This is because in later iterations changes in the objective reflect adjustments needed to satisfy the nonlinear state equation, which is imposed as a constraint but may be violated in the earlier iterations. These adjustments sometimes actually increase the objective function value. The converged objective function value of $\hat{J} = 3485$ compares well with the expected value of $\overline{J} = 3584$, lying just over one standard deviation away. Since the objective function can be evaluated at negligible cost in the representer algorithm [9], this is an important validation step when actual data are assimilated.

D. Model Error and Adjoint Variables

The adjoint variable λ plays a key role in our variational assimilation algorithm since it relates the model error and initial condition parameter estimates to the measurements. This is apparent from (6), which indicates that the model error estimate is the convolution of the adjoint with the model error covariance, and the parameter estimate is a weighted sum of the elements of the adjoint vector at the initial time. If the model errors were uncorrelated in time and space, the model error estimate for each component of the state equation would be equal to the corresponding adjoint variable. In our soil moisture application, however, temporal and spatial correlations are important since they account for persistence of uncertain inputs such as errors in precipitation and soil parameters.

An examination of the adjoint values obtained from our synthetic experiment provides some insight into the relationship between the adjoint variables and the model error. Fig. 5 compares time series of the soil moisture and soil temperature adjoint variables in a representative pixel with brightness temperature measurements and estimates in the same pixel. The soil moisture adjoint variables are provided at several different depths. The abrupt changes in the surface values of the adjoint variables observed as time progresses backward reflect the forcing of the backward (adjoint) (5) with impulses at the observation times. The weights of these impulses are proportional to the misfits between the observed and the estimated brightness temperatures and to the sensitivity derivative of the brightness temperature with respect to the state vector. The magnitude of the adjoint increases rapidly (moving backward in time) when the effects of a new measurement are felt and then gradually decays as these effects are dampened by the model dynamics.

The brightness temperature depends directly on the saturation in the top two nodes (at 0 cm and 5 cm) via the top layer microwave emissivity and the heat capacity. Consequently, the brightness temperature data misfit has an instantaneous effect on the surface soil moisture adjoint variable. Since saturation values in the lower layers do not directly influence the brightness temperature, the soil moisture, adjoint variables for these layers change more gradually, reflecting the delayed propagation of information downward from the surface. The soil moisture at all depths has a longer memory than the soil temperature, which decays (backward) to zero within 12 h after a measurement.

Fig. 6 plots estimated, prior, and true model error time series for the same pixel considered in Fig. 5. Note that the prior (or mean) model error is zero by convention. The upper panel shows the model error for the moisture flux upper boundary condition (the forcing term of the soil moisture balance equation). While the true model error correlation time is 10 h, we only assimilate brightness updates once per day so the model error estimate is necessarily much smoother than the true time series. Nevertheless, the estimate tends to track the general trend of the actual model error.

The second panel of Fig. 6 shows model errors in the soil temperature equation (forcing term for the soil energy balance). The estimate of the soil temperature model error is somewhat poorer than the estimate of the soil moisture model error. This observation applies generally, across the model domain. The third panel shows the model error in the canopy temperature equation (forcing term for the canopy energy balance). Apparently, we have no skill in estimating this error. The poor estimates obtained for the soil and canopy temperatures reflect the fact that
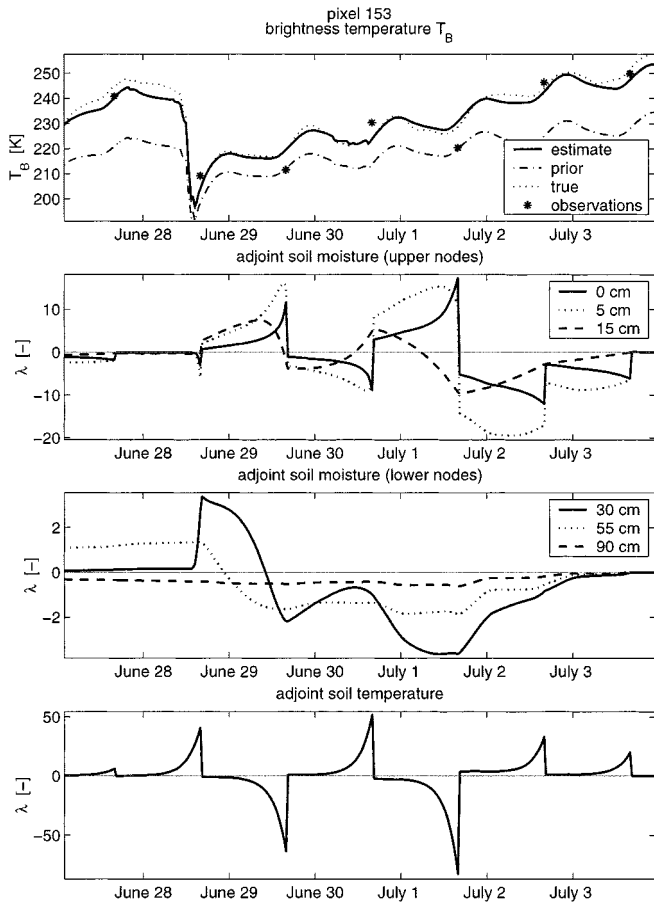
Fig. 5. Top panel shows the estimated brightness temperature for a representative pixel along with the (synthetic) observations. For reference, the prior and the true brightness are also shown. The second and third panels show the adjoint soil moisture for the upper and lower three nodes, respectively. The bottom panel shows the adjoint soil temperature.

Fig. 6. Model error (top) in the moisture flux upper boundary condition, (middle) in the soil energey balance, and (bottom) in the canopy energy balance for a representative pixel (the same as Fig. 5).

daily L-band brightness images contain little or no information on soil and canopy temperature at the 30 min scale. We are able to estimate the soil moisture model error because soil moisture has a longer memory than the soil or canopy temperatures. This is evident from the adjoint variable plots of Fig. 5.

E. Computational Effort

The computational burden for advanced data assimilation in operational land surface applications is formidable. This reflects the large horizontal extent of the model domain, the need for high resolution in time, the large number of remote sensing measurements to be assimilated, and the nonlinear nature of the state and measurement equations. It is useful to review some of the factors which affect the computational time. Such a review provides some insight about the limitations of current data assimilation methodology and suggests where improvements are most needed.

In the representer method, the bulk of the computational effort is spent on model integrations. If we count the forward and backward integrations separately and equally, the integration count for the direct iterated representer method is $(2N_Z + 3) = 7171$ model integrations per iteration [9]. For the indirect iterated representer method, the integration count depends on the number of linear combinations of representer fields needed in
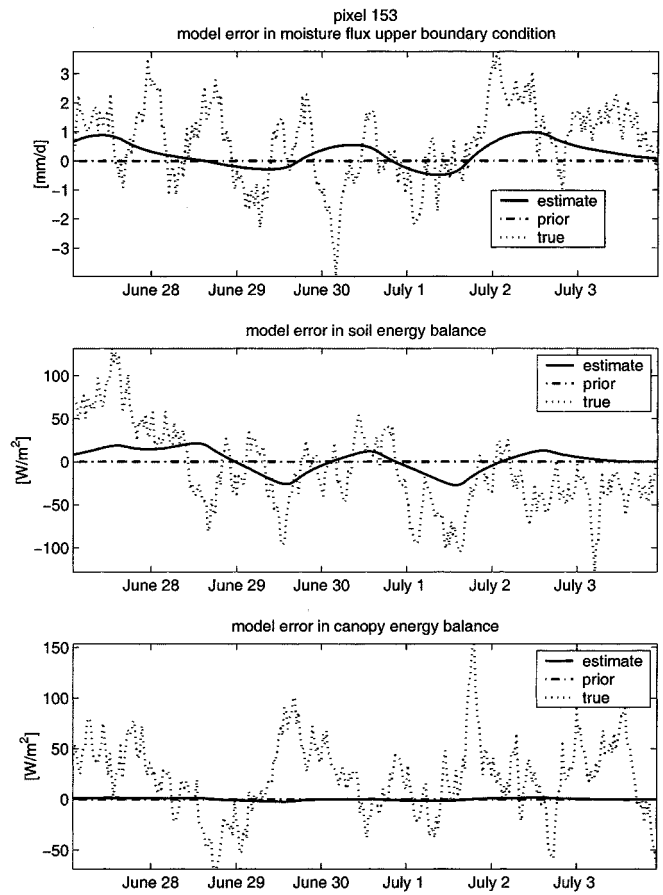
the conjugate gradient solver of (17). Table II shows the number of model integrations for each iteration of our synthetic test problem. For our example, the computational burden of the indirect iterated representer method is one order of magnitude less than that of the direct method. This efficiency has an intuitive physical interpretation. Obviously, the horizontal correlations present in the initial condition and model errors imply that there are many fewer degrees of freedom in the brightness images than there are scalar data. Through its design, the indirect method effectively compresses the data.

For operational applications, it is important to understand how computational requirements scale as we increase the size of the problem. Since the model is composed of a collection of independent vertical columns, the computational requirements for each model integration scale with the number of pixels N_p (for a constant number of layers in the vertical). We have found that the number of model integrations required in the indirect iterated representer method is mostly determined by the number of scalar data being assimilated [7]. Therefore, the computational effort scales roughly with the product of the number of measurements and the number of pixels $N_Z N_p \propto N_p^2$.

The memory requirements for the assimilation algorithm are substantial and scale with the length of the assimilation interval and the size of the grid. Ideally, all variables should be kept in memory during the assimilation to keep the cost of input/output operations low. In addition to the model inputs we have to store

the state trajectory over the entire interval $[0, t_f]$, and we need to store the estimates of the model error over $[0, t_f]$.

Finally, we briefly compare the computational requirements of the indirect iterated representer technique, the gradient-descent ("adjoint") method [9], the extended Kalman filter (EKF) [11], the ensemble Kalman filter (EnKF) [21], and error subspace statistical estimation (ESSE) [22]. Of these techniques, the representer method, the gradient-descent approach, and the EKF fully account for the temporal propagation of uncertainty. The EnKF and ESSE approximate the dynamic evolution of the error covariances. Our comparison is based on an approximate count of floating point operations rather than actual computational experience (see [7], [22] for details).

The EKF explicitly propagates the state error covariance matrices $\left(\propto N_p^3 \right)$ and cannot be used for 4-D large-scale soil moisture assimilation unless all horizontal correlations are neglected. Likewise, the adjoint-based gradient-descent method is not feasible if model errors are included because the number of degrees of freedom increases in direct proportion to the number of time steps. From a computational point of view, the indirect representer technique is competitive with approximate sequential methods like the EnKF and ESSE, which scale with $N_e N_p$, where N_e is the number of ensemble members. It is not yet well understood how many ensemble members are required, but we expect N_e to scale with N_p, the size of the domain. In this case, the indirect representer and Monte Carlo methods appear to be competitive, both requiring on the order of N_p^2 model integrations for large problems. For comparable effort, the great advantage of the representer technique lies with the near-optimality of the estimates, while the sequential Monte Carlo methods may be a better choice if posterior state covariance information is required.

VII. Summary and Conclusions

In this article, we have described a variational method which accounts for process dynamics and model errors and is suitable for large-scale soil moisture assimilation. We have applied the indirect iterated representer algorithm to a synthetic test experiment and demonstrated its near-optimal performance. This was done by closely examining the statistical properties of the posterior data residuals and by looking at the value of the objective function. Although nonlinearities in the model and in the measurement process complicate the performance assessment, our results support the conclusion that large-scale soil moisture data assimilation is feasible. Since similar validation techniques can also be applied in field tests, it should be possible to verify whether the approach works as well in practice as it does in our synthetic example.

We are confident that soil moisture can be satisfactorily estimated from daily L-band brightness temperature observations. However, it appears that soil and canopy temperatures cannot be adequately retrieved from daily observations alone. An examination of the adjoint variables reveals the shorter memory of the soil and canopy temperatures as compared to soil moisture. Other types of measurements, such as infrared remote sensing observations, are available more frequently and could be assimilated to improve the temperature estimates. An additional insight provided by the adjoint variables is the time lag inherent in propagating information from the surface down the soil profile.

Theoretical error covariances indicate that uncertainty in brightness temperature estimates (and therefore in soil moisture estimates) decreases with increasing soil moisture. At one extreme, we know that surface soil moisture is guaranteed to be close to saturation after a significant rainstorm. As precipitation inputs decrease and the effects of drydown dominate, uncertainty gradually increases. If drydown persists for a long time, soil moisture approaches residual saturation and the uncertainty decreases again.

Much more work is necessary before soil moisture data assimilation becomes practical in an operational setting. Obvious future research directions are a test of the algorithm with field data (for example, ESTAR radiobrightness observations collected during SGP97 [19]), and the assimilation of additional types of measurements. Moreover, operational applications require a scheme for reinitializing subsequent assimilation intervals. Since the necessary posterior state covariances are too expensive to compute in practice, this must be done in an approximate fashion. Preliminary results are encouraging [14].

In spite of the substantial gains in computational efficiency provided by the indirect iterated representer method, the computational effort is tremendous. Of particular concern is the fact that the computational load scales with the number of observations that are assimilated. We believe that operational applications must rely on parallel computing, which is quite compatible with the multicolumn structure of the hydrologic model used in our soil moisture application. Even if the computational limitations turn out to be prohibitive for operational applications in the near future, the proposed algorithm provides a valuable benchmark against which computationally more efficient algorithms that do not account for error covariance propagation can be evaluated.

References

[1] M. B. Parlange, G. G. Katul, M. V. Folegatti, and D. R. Nielsen, "Evaporation and the field scale soil water diffusivity function," *Water Resources Res.*, vol. 29, no. 4, pp. 1279–1286, 1993.

[2] D. Entekhabi, H. Nakamura, and E. G. Njoku, "Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 438–448, Mar. 1994.

[3] J.-C. Calvet, J. Noilhan, and P. Bessemoulin, "Retrieving the root-zone soil moisture from surface soil moisture or temperature estimates: A feasibility study based on field measurements," *J. Appl. Meteorol.*, vol. 37, no. 4, pp. 371–386, 1998.

[4] F. Castelli, D. Entekhabi, and E. Caporali, "Estimation of surface heat flux and an index of soil moisture using adjoint-state surface energy balance," *Water Resources Res.*, vol. 35, no. 10, pp. 3115–3125, 1999.

[5] J. F. Galantowicz, D. Entekhabi, and E. G. Njoku, "Tests of sequential data assimilation for retrieving profile soil moisture and temperature from observed L-band radiobrightness," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, p. 1860, July 1999.

[6] P. R. Houser, W. J. Shuttleworth, J. S. Famiglietti, H. V. Gupta, K. H. Syed, and D. C. Goodrich, "Integration of soil moisture remote sensing and hydrologic modeling using data assimilation," *Water Resources Res.*, vol. 34, no. 12, pp. 3405–3420, 1998.

[7] R. H. Reichle, "Variational assimilation of remote sensing data for land surface hydrologic applications," Ph.D. dissertation, Mass. Inst. Technol., Dept. Civil Environ. Eng., Cambridge, 2000.

[8] J.-N. Thépaut and P. Courtier, "Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model," *Q. J. R. Meteorol. Soc.*, vol. 117, pp. 1225–1254, 1991.

[9] A. F. Bennett, *Inverse Methods in Physical Oceanography*. New York: Cambridge Univ., 1992.

[10] P. Courtier, J. Derber, R. Errico, J.-F. Louis, and T. Vukícević, "Important literature on the use of adjoint, variational methods and the kalman filter in meteorology," *Tellus*, vol. 45A, pp. 342–357, 1993.

[11] A. Gelb, Ed., *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.

[12] A. F. Bennett, "Inverse methods and data assimilation," Lecture Notes, 1999 Summer School, College Ocean. Atmos. Sci., Oregon State Univ., Corvallis, OR, July 1999.

[13] A. F. Bennett, B. S. Chua, and L. M. Leslie, "Generalized inversion of a global numerical weather prediction model," *Meteorol. Atmos. Phys.*, vol. 60, pp. 165–178, 1996.

[14] R. H. Reichle, D. Entekhabi, and D. B. McLaughlin, "Downscaling of radiobrightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach," Water Resources Res., 2001, to be published.

[15] R. B. Clapp and G. M. Hornberger, "Empirical equations for some soil hydraulic properties," *Water Resources Res.*, vol. 14, pp. 601–604, 1978.

[16] Z. Hu and S. Islam, "Prediction of ground surface temperature and soil moisture content by the force-restore method," *Water Resources Res.*, vol. 31, pp. 2531–2539, 1995.

[17] Y. Xue, P. J. Sellers, J. L. Kinter, and J. Shukla, "A simplified biosphere model for global climate studies," *J. Climate*, vol. 4, pp. 345–364, 1991.

[18] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing*. Norwood, MA: Artech House, 1986.

[19] T. J. Jackson, D. M. Le Vine, A. Y. Hsu, A. Oldak, P. J. Starks, C. T. Swift, J. D. Isham, and M. Haken, "Soil moisture mapping at regional scales using microwave radiometry: The southern great plains hydrology experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2136–2151, Sept. 1999.

[20] J. R. Benjamin and C. A. Cornell, *Probability, Statistics and Decision for Civil Engineers*. New York: McGraw-Hill, 1970.

[21] G. Burgers, P. J. van Leeuwen, and G. Evensen, "Analysis scheme in the Ensemble Kalman filter," *Mon. Weather Rev.*, vol. 126, pp. 1719–1724, 1998.

[22] P. F. J. Lermusiaux and A. R. Robinson, "Data assimilation via error subspace statistical estimation. Part I: theory and schemes," *Mon. Weather Rev.*, vol. 127, pp. 1385–1407, 1999.

Rolf H. Reichle received the M.S. degree in physics (with honors) from the University of Heidelberg, Germany, in 1996, and the Ph.D. degree in environmental engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 2000.

He is currently a Member of the Research Faculty, University of Maryland, Baltimore County, Baltimore. He is currently with the NASA Goddard Space Flight Center, Greenbelt, MD, working on seasonal-to-interannual prediction of climate variability. His research interests include land–atmosphere interactions, data assimilation, and remote sensing.

Dennis B. McLaughlin received the Ph.D. degree from Princeton University, Princeton, NJ, in 1985.

He is currently the H. M. King Bhumibol Professor of Water Resources Management, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology (MIT), Cambridge. His research interests include data assimilation, stochastic solute transport, and management of water resources in semi-arid regions. He is a member of the American Geophysical Union and a Deputy Editor of *Water Resources Research*.

Dara Entekhabi received the Ph.D. degree from the Massachusetts Institute of Technology (MIT), Cambridge, in 1990.

He is a Professor with the Department of Civil and Environmental Engineering, MIT. His area of research focus includes terrestrial remote sensing, data assimilation, and coupled land-atmosphere systems behavior.

Dr. Entekhabi is Chair of the National Research Council Committee on Hydrologic Science, a member of the American Meteorological Society, and a Fellow of the American Geophysical Union.